



4184-09 - P

## **DEPARTMENT OF HEALTH AND HUMAN SERVICES**

### **Administration for Children and Families**

#### **Criteria for Evidence of Effectiveness to be Applied to Projects Identified for Inclusion in the What Works Clearinghouse of Proven and Promising Projects to Move Welfare Recipients into Work**

**AGENCY:** Administration for Children and Families (ACF), U.S. Department of Health and Human Services (HHS)

**ACTION:** Request for public comment.

**SUMMARY:** The Administration for Children and Families, HHS, solicits comments by August 5, 2018 on the criteria for evidence of effectiveness for the What Works Clearinghouse of Proven and Promising Projects to Move Welfare Recipients into Work. Final criteria for evidence of effectiveness will be used to develop the clearinghouse.

#### **SUPPLEMENTARY INFORMATION:**

*Invitation to Comment:* HHS invites comments regarding this notice on the proposed criteria for HHS's systematic review of the evidence. To ensure that your comments are clearly stated, please identify the specific criterion or other section of this notice that your comments address.

### **1.0 Background**

#### **1.1 Legislative Context**

The Consolidated Appropriations Act of 2017 (Pub. L. 115-31 (<https://www.congress.gov/115/plaws/publ31/PLAW-115publ31.pdf>)) directs the U.S. Department of Health and Human Services (HHS) to create a database of projects that

have used a proven or promising approach to move welfare recipients into work, based on independent, rigorous evaluations of the projects, and to create a What Works Clearinghouse of Proven and Promising Projects to Move Welfare Recipients into Work. As stated in the statute, the database shall additionally “include a separate listing of projects that used a developmental approach in delivering services and a further separate listing of the projects with no or negative effects.” The statute requires HHS to establish criteria for evidence of effectiveness.

#### 1.2 The Legislation’s Direction for Establishing the Criteria for Evidence of Effectiveness

Section 413(g)(2) of Pub. L. 115-31 charges the Secretary of Health and Human Services with establishing the criteria of effectiveness. The statute further stipulated that the (B) process for establishing the criteria –

- (i) is transparent;
- (ii) is consistent across agencies;
- (iii) provides opportunity for public comment; and
- (iv) takes into account efforts of Federal agencies to identify and

publicize

effective interventions, including efforts at the Department of Health

and

Human Services, the Department of Education, and the Department of Justice.

#### 1.3 The Employment Strategies for Low-Income Adults Evidence Review

Prior to the enactment of Pub. L. 115-31, the Office of Planning, Research, and Evaluation (OPRE) at the Administration for Children and Families (ACF) at HHS had developed the Employment Strategies for Low-Income Adults Evidence Review (ESER). The new statute aligns with and extends the work of ESER. HHS proposes building on this existing work to develop the new Clearinghouse.

The Employment Strategies for Low-Income Adults Evidence Review (ESER) is a systematic review of the evaluation research published between 1990 and 2014 on employment and training programs for low-income adults. It culminated in a searchable, public database (<https://employmentstrategies.acf.hhs.gov/>). The review was supplemented with briefs synthesizing the results of the review and highlighting strategies that appeared to be promising, as identified by the review. To identify the programs and strategies--or interventions-- that appear to be most effective in helping low-income adults gain and retain employment, ESER systematically identified, assessed, and synthesized evidence from the existing evaluation research literature. A core component of ESER's review, as with other federal evidence reviews, involved assessing the quality of the research evidence on different interventions.

To assess the quality of the evidence, ESER reviewed each study's methods to determine if they were rigorous enough to ensure that the study's findings could be considered reliable. ESER assessed whether the study's methods reliably supported the conclusion that an intervention's impacts were caused by the intervention and not by something else. The standards for assessing studies' methods were defined based on consultation with federal experts on evidence reviews and researchers with expertise in evaluation methodology. To differentiate among different levels of the strength of evidence, ESER

assigned a High, Moderate, or Low rating to each study reviewed. 246 of the 314 studies included in the review earned a High rating and 1 study earned a Moderate rating. The remaining 67 studies received a Low rating.

Through this review, ESER was able to identify interventions whose findings could be considered most reliable. Studies' ratings reflect the rigor of their study methods, independent of whether the findings were positive or negative. As a result, a study could be rated High or Moderate even if the intervention studied did not improve the outcomes for low-income adults. While the vast majority of studies included in ESER achieved a High rating (and, therefore, are considered to provide reliable, or strong, evidence), the review also found that, overall, null impacts were more prevalent than statistically significant impacts.

While ESER did not assess the effectiveness of the interventions reviewed, ESER conducted a number of preliminary steps necessary for assessing effectiveness. This included categorizing each study's findings according to whether it found positive, negative or null impacts for the interventions studied. In addition, through a number of synthesis briefs (published on the website), ESER qualitatively and quantitatively summarized the direction of impacts for different interventions and highlighted interventions associated with the greatest number of positive impacts.

To be included in ESER, studies had to -

- Quantitatively measure the effectiveness of a program or strategy
- Be published between 1990 and 2014
- Study an employment program or strategy-- an intervention-- that

- had a primary aim of improving employment-related outcomes
- primarily targeted low-income adults
- took place in the United States, Canada, or the United Kingdom

To identify studies eligible for review, ESER issued a call for papers, conducted literature searches, and consulted with experts in workforce development programs that serve low-income adults.

ESER looked at the effects of the interventions on four domains, or outcome areas:

- Employment
- Earnings
- Public benefit receipt
- Education/training

Outcomes were examined for short and longer-term impacts (longer-term was measured as being more than 18 months after the intervention was implemented).

The ESER website (<https://employmentstrategies.acf.hhs.gov/>) reports key results for all eligible studies. The website also allows users to search for results by program studied, target population, outcome(s) of interest, service strategies, intervention setting, year of study publication, and whether favorable impacts were found.

While ESER's overall population of interest was low-income adults, a majority of the studies in ESER examined welfare populations. Because studies of interventions in a welfare setting typically include both recipients and applicants, ESER does not include any studies that solely focused on welfare recipients. ESER does, however, include interventions targeted to low-income populations understood to share important

characteristics with welfare recipients, such as other public benefit recipients, and those considered hard to employ, including those who have been homeless or formerly incarcerated.

## 2.0 Process for Establishing the Criteria of Effectiveness for the new What Works Clearinghouse

In fall 2017 and early winter 2018, OPRE engaged in a series of systematic consultations with federal and non-federal technical experts on evidence reviews. In addition to representation from the Department of Labor (DOL) and the Department of Education (ED) in these consultations, federal representation included the Department of Justice (DOJ) and a number of HHS agencies/offices including the Office of Family Assistance (OFA), the Office of Planning, Research, and Evaluation (OPRE), the Office of the Assistant Secretary for Planning and Evaluation (ASPE) and the Agency for Healthcare Research and Quality (AHRQ).

The objective of these consultations was to help HHS:

- 1) develop criteria for categorizing interventions in the new Clearinghouse as proven, promising, developmental, or ineffective,
- 2) develop these criteria through a process that
  - a. involved consultation with the Department of Labor (DOL), the Department of Education (ED), and other entities with experience evaluating relevant effectiveness research,
  - b. allowed HHS to better understand other Federal evidence reviews' standards and processes and determine where it would make sense for the

new Clearinghouse to be consistent with these standards and processes,  
and

- 3) learn best practices from other Federal evidence reviews for identifying and publicizing effective interventions

## 2.1 Transparent

To ensure that the Clearinghouse's procedures and standards, including the criteria for evidence of effectiveness, are transparent, HHS intends to implement the following practices:

- Post the procedures and standards and information about the process on the Clearinghouse website.
- Provide the public a means of contacting the Clearinghouse, for example, by establishing a help desk to respond to e-mail inquiries.

## 2.2 Consistent Across Agencies

To ensure that the Clearinghouse is as consistent as possible with other federal evidence reviews in its processes and standards, HHS intends to implement the following practices:

- Adopt the standards and methods for reviewing studies from OPRE's existing Employment Strategies Evidence Review (ESER) (<https://www.acf.hhs.gov/opre/resource/employment-strategies-for-low-income-adults-evidence-review-standards-and-methods>), which are broadly consistent with other federal Clearinghouses. ESER's standards and methods (e.g., author queries;

number and training of reviewers; choices about reporting effect sizes) were developed by considering both the choices made by other federal and non-federal Clearinghouses and the standards of research in the employment and training intervention field. Other existing federal Clearinghouses have followed this same approach (considering both the choices made by other clearinghouses and the norms of research within their fields of focus).

- In any instances where the new Clearinghouse's ratings of a project's strength of evidence or effectiveness differ from another federal evidence review that rates projects according to the same outcomes (such as the Department of Labor's (DOL's) Clearinghouse for Labor Evaluation and Research (CLEAR)), annotate the findings to explain the reason for the difference.

### 2.3 Provides Opportunity for Public Comment

To provide an opportunity for public comment on the criteria for effectiveness, ACF is publishing this Federal Register Notice.

### 2.4 Takes into Account Efforts of Federal Agencies to Identify and Publicize Effective Interventions

To ensure the Clearinghouse reflects the learning of other Federal agencies about how to identify and publicize effective interventions, HHS intends to implement the following practices:



- Use some of the methods adopted by other clearinghouses to create multiple products tailored to different audiences and use graphic design and other user-friendly dissemination elements to help users digest evidence quickly.
- Include information on the Clearinghouse website that is especially useful to practitioners, such as summary information about projects and approaches.
- Develop and incorporate alternative media for the Clearinghouse such as videos that will tailor communication to various groups.
- Ensure that information is effectively conveyed on the Clearinghouse website by soliciting feedback from various stakeholders who can represent key target audiences. Key among these would be state or county Temporary Assistance for Needy Families (TANF) and Workforce Development practitioners, as well as evaluation researchers.

### 3.0 Proposed Criteria for Evidence of Effectiveness

#### 3.1 Criteria for Well-Designed, Rigorous Impact Research

HHS intends to employ the criteria established by OPRE's Employment Strategies for Low-Income Adults Evidence Review (ESER) to assess the quality of study design and to assess the strength of the evidence resulting from studies. These criteria (referred to as "standards and methods") are available in ESER's Standards and Methods report <https://www.acf.hhs.gov/opre/resource/employment-strategies-for-low-income-adults-evidence-review-standards-and-methods>

## 3.2 Proposed Criteria for Evidence of Effectiveness for Projects Included in the Clearinghouse

### 3.2.1 Definition of Project and Approach

The legislation requires that ratings, or categorizations, of evidence of effectiveness be applied to projects and approaches. To standardize definitions for these terms, HHS intends to define a project and an approach as follows:

- Define project to be a specific bundle of services and/or policies implemented in a given context.
- Project will be the unit that receives an effectiveness rating (i.e. proven, promising, developmental, or ineffective).
- Define approach to be the guiding framework of specific services (e.g., career pathways)
- Approaches will not be rated as proven, promising, developmental, or ineffective, but the Clearinghouse will include narrative summaries related to different approaches.
- While the legislation does not require HHS to define or evaluate the effectiveness of program components, there is interest in the field in examining program components. Thus, HHS intends that the Clearinghouse include meta-analyses of specific components of projects (such as “case management” or “job search assistance”) whenever appropriate and feasible.

### 3.2.2 Parameters Guiding the Application of Evidence of Effectiveness Ratings

Before a project can be categorized as being proven, promising, developmental, or ineffective, a number of preliminary definitions, or parameters, must be established to guide decision making. These include the outcomes for which a project's effectiveness will be evaluated, how a favorable or unfavorable effect will be measured, and how an effectiveness rating will be applied to a project.

#### 3.2.2.1 Outcomes

HHS intends that the new Clearinghouse will review the following outcomes:

- employment (short and longer-term),
- earnings (short and longer-term),
- educational attainment, and
- public benefit receipt.

#### 3.2.2.2 Definition of Favorable and Unfavorable Effects

HHS intends that the Clearinghouse consider only statistically significant findings ( $p < .05$ ) as evidence of favorable or unfavorable effects.

#### 3.2.2.3 Pre-defining criteria for selecting among multiple outcome measures

HHS intends to reduce the likelihood for reporting a false positive rate for outcomes – an issue that can occur when studies use multiple measures or multiple outcomes to assess impacts in the same domain (e.g., short-term earnings) – by relying on the decision rules ESER developed to address the potential for multiple comparisons.

<https://www.acf.hhs.gov/opre/resource/employment-strategies-for-low-income-adults-evidence-review-standards-and-methods>).

#### 3.2.2.4 Application of Evidence of Effectiveness Ratings

HHS intends that evidence of effectiveness ratings will be applied within outcome domains; each project will receive ratings of effectiveness on *each* outcome domain (e.g., a project may be found promising for short-term employment but ineffective for long-term employment). There will be *no* overall rating for the project.

#### 3.2.3 Definition of Proven

The legislation directs HHS to categorize projects as Proven, Promising, Developmental, or Ineffective.

HHS intends that for a project to be considered proven, the following conditions must be met:

- There must be at least two separate studies of the same project that meet evidence standards and meet criteria for a promising rating.
  - Studies are considered to be separate studies of the same project if they use non-overlapping samples to examine distinct implementations of the project.
- There must be only favorable or null impacts within a given outcome domain. Thus, no studies that meet evidence standards for a given outcome domain can show an unfavorable impact within that domain.
- Projects that have both favorable and unfavorable impacts in a given domain will be categorized as mixed.
- A project has a limited number, or proportion, of null findings in a given domain.

HHS is soliciting comments on how to best determine the ceiling for the number, or proportion, of null to positive findings in a given domain.

If subsequent studies or replications result in only null findings in a given domain, the review will establish procedures for revisiting a project's rating of proven.

#### 3.2.4 Definition of Promising

HHS intends that for a project to be considered promising, the following conditions must be met:

- One study of a project must meet evidence standards.
- That study must find only favorable or null impacts within a given outcome domain. Thus no studies that meet evidence standards for an outcome domain can show an unfavorable impact within the domain.
  - If the review examines more than one measure to identify impacts on a particular domain (e.g., Unemployment Insurance data and participant survey data), as long as one measure (among those selected according to 3.2.2.3 above) finds favorable impacts for that outcome, the intervention can receive a Promising rating for that outcome.
- Projects that have both favorable and unfavorable impacts in a given domain will be categorized as mixed.

#### 3.2.5 Definition of Ineffective

HHS intends that for a project to be considered ineffective, the following conditions must be met:

- One or more studies of a project must meet evidence standards.
- There must be only findings of unfavorable or null effects in a given domain.
- For studies finding null effect in a given domain, the review will include a measure of statistical precision -- so that small, under-powered studies do not drive the effectiveness rating. If an intervention has been evaluated using only small studies, a lack of detectable effects could reflect either ineffectiveness of the intervention or the lack of statistical power to detect effects. It would be misleading to characterize this latter scenario as an ineffective project.

### 3.2.6 Definition of Developmental

HHS intends that for a project to be considered developmental, the following conditions must be met:

- There must be at least one current, ongoing evaluation of the project that uses a study design that meets evidence standards but has not yet produced impact findings.

### 3.2.7 Additional Category of Mixed and Definition of Mixed

HHS intends that there be an additional category for categorizing evidence of effectiveness called mixed. HHS proposes that for a project to be considered mixed, the following conditions must be met:

- One or more studies of a project must meet evidence standards.
- The studies find both favorable and unfavorable impact estimates within the same domain.

3.2.8 HHS intends that narrative descriptions of rated projects, narrative descriptions of approaches, and information on case studies be provided to users of the Clearinghouse to facilitate a fuller understanding of the field of welfare-to-work interventions.

#### 4.0 Submission of Comments

Comments may be submitted until August 5, 2018 by email to  
OPREinfocollection@acf.hhs.gov.

**Naomi Goldstein,**

*Deputy Assistant Secretary for Planning, Research, and Evaluation.*

[FR Doc. 2018-12160 Filed: 6/5/2018 8:45 am; Publication Date: 6/6/2018]